ADAPTING SCALE FOR CHILDREN: A PRACTICAL MODEL FOR RESEARCHERS

Selami AYDIN[1], Leyla HARPUTLU[2], Şeyda Savran ÇELİK[3], Özgehan UŞTUK[4], Serhat GÜZEL[5], Deniz GENÇ[6]

[1, 4, 5, 6] Balikesir University, [2] Dokuz Eylül University, [3] Hakkari University
saydin@balikesir.edu.tr, leylaharputlu@gmail.com, seydasavran@balikesir.edu.tr, oustuk@balikesir.edu.tr, serhatguzel@balikesir.edu.tr, deniz-genc@hotmail.com

ABSTRACT

Measurement of children's behaviors in an educational and research context is a problematic and complex area. It is also evident that adapting scales to measure children's behaviors in an educational and research context is a complex process due to several reasons. First, cultural elements constitute a considerable problem. Second, it is difficult to obtain reliable and valid data in various educational and research settings. Third, language can be a significant variable, as children may have a low proficient level of the target language. Thus, the current paper aims to report on developing a practical model for researchers who need to adapt scales from one language to another, when measuring children's behaviors in an educational and research context. The model consists of four steps: First, how scales should be translated from one language to another is clarified. Then, the paper focuses on simplification and moderation. Then, how the pilot study should be administered is explained. Finally, how to validate scales are mentioned. The study concludes that the model proposed may be beneficial for researchers who need to adapt scales for children that were designed for adult learners and in a different language. It is also concluded that reliability and validity of the data collected from children are strictly related to the steps listed in the study.

Key Words: Scale, Adaptation, Validation, Children.

ÖZET

Eğitim ve araştırma bağlamında çocukların davranışlarının ölçülmesi her zaman zorlu ve karmaşık olarak karşımıza çıkmaktadır. Farklı nedenlere dayalı olarak, eğitim ve araştırma bağlamlarında çocukların davranışlarının yönelik olarak, mevcut ölçeklerin adapte edilmesi de ayrıca karmaşık bir süreçtir. Öncelikle, olarak kültürel ögeler, dikkate değer bir problem alanıdır. İkinci olarak, çeşitli eğitim ve araştırma ortamlarında güvenilir ve geçerli veri elde edebilmenin zor bir süreç olduğu açıktır. Üçüncü olarak, çocukların hedef dildeki düşük dil yeterliliklerinden ötürü, dilin bizzat kendisi çok önemli bir değişken olabilmektedir. Bu sebeplere dayanarak, mevcut çalışma, araştırmacılara, eğitim ve araştırma bağlamında çocuklardan veri toplarken ölçekleri bir dilden diğerine çevirmek durumunda uygulanabilir bir model önermektedir. Bu model dört adımdan oluşmaktadır. İlk adımda ölçeklerin ne şekilde tercüme edilmesi gerektiği açıklanmıştır. İkinci adımda, ölçeğin sadeleştirilmesi ve uyarlanması konularına değinilmektedir. Üçüncü adımda, pilot çalışmanın ne şekilde uygulanması gerektiği açıklanmış ve son olarak ölçeğin geçerlilik doğrulanmasının ne şekilde yapılması gerektiğinden bahsedilmektedir. Çalışmanın sonunda, dört aşamalı bu modelin, yetişkinler için tasarlanıp oluşturulan anket ve ölçekleri çocuklara uyarlamaya gereksinim duyan araştırmacılar için yararlı olabileceğinin altı çizilmiştir. Sonuç olarak, çocuklardan toplanan verinin güvenilirliği ve geçerliliği çalışmada listelenen adımlarla doğrudan ilişkilidir.

Anahtar Sözcükler: Ölçek, Uyarlama, Doğrulama, Çocuklar.

Introduction

When measuring children's behaviors in a research context, there are several features that should be considered. First, children have a short attention span and enjoy plays and stories. Second, children like pictures, colors and attractive typography. Third, given that their linguistic and cognitive abilities are still developing, testing and measurement should be performed in their native language. Fourth, as

children learn easier through social interaction, tasks should involve peer and group work. Last, integrative tests should be used, as they need to use two or more senses and skills (Hughes, 2003). In conclusion, it is evident that measuring children's behaviors requires a special attention in the research context.

The scales adapted for collecting data are supposed to provide reliable and valid information that is reflected explicitly and objectively. However, the use of the same research scales of different languages and cultures brings about several disadvantages. Whilst conducting a scale, it should be realized that cultural differences inherently exist between the original source of the scale and the target context in which it will be carried out. Owing to questions and items having dissimilar meanings in terms of different cultural features and values, the validity and reliability of the scale are under threat (Rode, 2005). Namely, the quality of scales that was from different cultural backgrounds and consisting of non-validated translations leads to incorrect results that do not reflect valid and reliable data (Auer et. al., 2000). Scales in a foreign language including specific characteristics of a culture that are not overlapped by the target culture are shown to be ineffective by researchers. For the purpose of obtaining proper data, the instruments need to be in native language of the participants, culturally acceptable or appropriately translated to be statistically valid (Cha, et. al. 2007). As a concluding point, although few studies have adapted scales including scale translation and validation for other cultures (Auer et. al., 2000), the problem still remains as an unresolved issue.

It is acknowledged that the amount of scales developed specifically for children is quite insufficient for it is hard to detect a framework on how to design appropriate scales for children. Although it is quite possible to detect scales developed for children in certain areas such as medicine, clinical psychology, it is not very common to trace a well-developed scale for children in social sciences. As the reason for the lack of scales, Parker (1984) suggests that children are much more complicated to gather information when compared to adults and nature of their constant development hinders consistency. Parker (1984) also states that children experience four stages of development as: preschool, middle school, early adolescence and middle adolescence. Therefore, these phases bring about dramatic changes to children's lives, and require different approaches, when it comes to gathering data from them. In addition, dealing with children is a matter of sensitivity, since they are quite fragile in terms of cognitive and psychological issues (Mahon et al., 1996). Moreover, as Vaillancourt (1973) claims, children's responses are not stable due to certain reasons such as the researcher's presence, the time interval between scale administrations and most importantly, item-related problems. To be clear, some items are quite hard for children to understand; that is why, they need to be simplified and adjusted to their levels.

In conclusion, several problems guided this study. First, in terms of scale adaptation, as children differ from adults in their psychological, cognitive and social characteristics, children's psychological, cognitive and social attention span, linguistic skills and social tendencies should be taken into account. Thus, the scales that are used to measure adults' behavior samples should be adapted for children. Second, when researchers prefer administering a scale in another language, cultural differences may be a factor that causes distortion of the data collected. For this reason, scales are needed to be translated for the utilization for participants from different cultures. Given that the mentioned issues may constitute problems during the measurement of children's behaviors, a strict procedure should be followed in terms of validity and reliability of a scale. Thus, this paper aims to present a practical model for researchers who need to translate and adapt scales for children.

The Model

The model proposes four steps: (1) Translation, (2) simplification and moderation, (3) piloting and (4) validation.

Translation

For the translation process, it is recommended that at least five translators are necessary. Among those translators, one should be an expert in the related field. In other words, the expert should be at least an MA or Ph.D. degree, as the process requires research experience in the related research field. Another one should be an expert in the field of child development, whereas a third one should be fully and culturally proficient in both languages and cultures. The fourth one is a translator who is independent from both the related field and child development and have a moderate language level in the target language. The last one should be a drama expert who is experienced in drama activities with children and be proficient in both languages.

After fully informing the translators about their roles in the procedure, they can translate the scale from the target language to the native language in a blind session. By this way, it is possible to reach five independent versions of the scale. In the following step, they need to compare their individual versions and unify them into one draft. For this purpose, panels should be repeated until a satisfactory equivalence version is reached. During the panels, all the translators should focus on the conceptual and semantic equivalence in order to ensure ideas, notions and meaning that do not exist in any of the languages.

After designing the translated version of the scale, it is necessary to administer both the original and translated scales to adult learners. Now then, it is possible to copy and print both versions and number them before the administration. At this point, it is recommended that first the original version should be administered. After a moderate time that aims to decrease the time effect on participants' behaviors, the translated version can be administered.

Following step includes the statistical analysis of the data for a deeper understanding of the similarities and differences between the two scales. For this purpose, Cronbach's Alpha can be calculated to see the extent to which the items in the two scales represent internal consistency. Then, an explanatory factor analysis should be carried out to compare the relationships between the items in both scales. In addition, a principal component analysis should be performed to see the extent the scales reflect the construct validity within the same sample group. Finally, the Varimax method and two-factor solution can be used to understand whether the equivalence between the scales is established. To conclude, after seeing that two scales have internal consistency and equivalence in terms of construct validity, and that both versions show a near identical factor loadings structure, it is possible to moderate the new version of the scale.

Simplification and moderation

The scale obtained after the translation process is appropriate for adults. Thus, it is necessary to simplify and moderate it for children. In other words, each item in the translated scale is needed to be simplified in accordance with linguistic and conceptual developments of children. For this purpose, the panel members, consisting of an expert in the field, an expert in the field of child development, a linguist who is proficient in children's native language, an independent one who has a moderate level language in the native language and a drama expert who is experienced in working with children should simplify the items in the scale in a blind session. In this process, it is expected that all members should pay attention

to children's conceptual and linguistic developments. Then, in panels, members should discuss each item and reach a consensus regarding respondency and intelligibility among children.

Then, the scale simplified should be moderated to see how children perceive each item in the scale. For this purpose, interactional role plays, process drama strategies, individual, peer and group drama activities can be performed. However, it should be underlined that the mentioned process should be carried out by a drama expert who is experienced in working with children. All of the performance should be recorded and saved for panel investigation. After the performance of the process, in a new panel, all of the panelists should examine the audio and visual recordings to see children's reactions to each of the items. Then, after reaching a consensus in terms of the respondency and intelligibility among children, the items in the scale can be rewritten and restructured for the administration process.

When restructuring the scale, several points should be considered. First, a clear and concise instruction is necessary for the administrator. This section may vary in accordance with the techniques that are used to collect data.

Figure 1: Sample Instruction Page

To the administrator;
This scale includes three demographic questions and 20 items related to foreign language anxiety among young learners.

Each item has a statement and facial images. The first face is **very unhappy**; the second is **unhappy**; the third is **neither happy nor unhappy**; the fourth one is **happy**; the last one is **very happy**.
Please ask students to write their grade, gender and age. Then, ask them to read and choose the most appropriate facial expression for each statement. Please read aloud the statements if necessary.

Then, demographic variables such as age and gender should be asked directly in a background questionnaire, as shown below. It should be noted that these factors may vary in accordance with the research context and content. As a note, instead of asking children's ages, their birthday should be interrogated. In addition, "boy or girl" can be preferred instead of "female or male".

Figure 2: Sample Background Questionnaire

What is your grade?   : _____

Are you a boy or a girl   : _____

When is your birthday? : _____

When designing the items in a Likert scale, statements such as "strongly disagree, disagree, neither agree nor disagree, agree, strongly agree, never, rarely, sometimes, usually, always" should be avoided. Instead, emoticons, colors or pictures should be preferred, as shown below.

Figure 3. Sample item

1. When I speak in my English class, I feel:
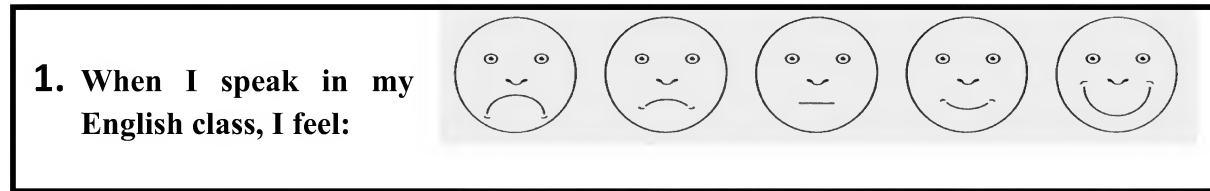
Piloting

Similar to the moderation process, the administration of piloting phase should be performed by the drama expert. For this purpose, several techniques such as interactional role plays, process drama strategies and individual, peer and group drama activities can be used. The process should start with a pre-text that initiates the dramatic world in which the implementation of scale may take place. This pre-text can be a word, a location, a gesture, an object, a story, an image or an idea (O'Neill, 1995). For instance, they may adopt the role of guest investigators to inquire a problem. The introduction of problem and scale should include pair and group works; however, because the scale should be answered individually, dramatic activities should evolve to individual works and end up with individual respondency to the scale. The reason behind this individuality is clarified to the children within the boundaries of the dramatic world. After the play is implemented, the follow-up explaining why such data are collected from them and letting them tell what their reflections should be made to end the dramatic activity properly.

The data obtained from the pilot testing need to be analyzed in terms of demographic information, reliability and validity. For this purpose, first, gender and age frequencies can be computed. To add, mean score for age can be found. Then, after the descriptive data are calculated in terms of mean scores and standard deviations related to the items, Cronbach's Alpha and Cronbach's Alpha Based on Standardized Items should be computed to see the extent to which the items represent reliability. Then, for construct validity, an exploratory factor analysis should be performed. For this purpose, a principal component analysis and the Varimax method should be carried out. In this analysis, principal components with Varimax rotation should be used. In addition, the eigenvalues, amount of variance explained and scree test should be performed to obtain an optimal factor solution. After this process, items that are not functioning or not related to any factor can be removed from the scale. Now then, the scale is ready for validation (Aydin et al. 2016a).

Validation

Validation process involves the same procedure with the piloting phase. However, in the process, it is necessary to use another and larger sample group. As previously underlined, this process should be also performed by the drama expert. The same techniques used to collect data during the piloting should be used in the validation process.

In this process, repeatedly, the data should be analyzed in terms of demographic information, reliability and validity. After computing age and gender frequencies, mean scores for age can be computed. Then, after mean scores and standard deviations related to the items derived from the pilot study are found, Cronbach's Alpha and Cronbach's Alpha Based on Standardized Items are calculated to see the extent

to which the items show reliability. After performing a principal component analysis and the Varimax method, the eigenvalues, amount of variance explained and scree test are performed to see the optimal factor solution. As conducted in piloting phase, unrelated and non-functional items should be removed. The version of the scale reached now then is ready for measuring children's behaviors (Aydin et al. 2016b).

Conclusion

The paper reports on developing a practical model for researchers who need to adapt scales from one language to another, when measuring children's behaviors in an educational and research context due to several reasons. Cultural elements, different educational settings and language barriers may distort data, whereas measurement of children's behaviors in an educational and research context is a problematic and complex area. Thus, the paper proposes a four-step model to adapt scales to administer to children. The steps involve translation, simplification and moderation, piloting and validation. It is expected that the model proposed is beneficial for researchers who is interested in adapting scales for children.

Acknowledgements

References

Auer, S., Hampel, H., Möller, H. J., & Reisberg, B. (2000). Translations of measurements and scales: Opportunities and diversities. *International Psychogeriatrics, 12*(1), 391-394.

Aydın, S., Harputlu, L., Güzel, S., Savran Çelik, Ş., Uştuk, Ö. & Genç, D. (2016a). A Turkish version of Foreign Language Anxiety Scale: Reliability and validity. Paper presented at GlobELT 2016 Conference: An International Conference on Teaching and Learning English as an Additional Language, 14 – 17 April, 2016, Antalya, Turkey.

Aydın, S., Harputlu, L., Güzel, S., Savran-Çelik, Ş., Uştuk, Ö. & Genç, D. (2016). Children's Foreign Language Anxiety Scale: Preliminary tests of reliability and validity. *Turkish Online Journal of English Language Teaching, 1*(3), 144-150.

Bolton, G. & Heathcote, D. (1999). *So you want to use role-play? A new approach in how to plan*. Staffordshire: Trentham Books.

Cha, E. S., Kim, K. H., & Erlen, J. A. (2007). Translation of scales in cross-cultural research: Issues and techniques. *Journal of Advanced Nursing, 58*(4), 386-395.

Hughes, A. (2003). *Testing for language teachers*. Cambridge University Press.

Mahon, A., Glendinning, C., Clarke, K., & Craig, G. (1996). Researching children: Methods and ethics. *Children & Society, 10*(2), 145-154.

O'Neill, C. (1999). *Process drama: A framework for process drama*. Portsmouth: Heinemann.

Parker, W. C. (1984). Interviewing children: problems and promise. *The Journal of Negro Education, 53*(1), 18-28.

Rode, N. (2005). Translation of measurement instruments and their reliability: An example of job-related affective well-being scale. *Advances in Methodology and Statistics, 2*(1), 15-26.

Vaillancourt, P. M. (1973). Stability of children's survey responses. *The Public Opinion Quarterly, 37*(3), 373-387.